AN OVERVIEW OF THE ETHERPHONE SYSTEM AND ITS APPLICATIONS*

Polle T. Zellweger, Douglas B. Terry, and Daniel C. Swinehart

P.C. 07/03/88 (9)
p. 160-168

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304

Abstract

The Etherphone™ system has been developed to explore methods for extending existing multi-media office environments with the facilities needed to handle the transmission, storage, and manipulation of voice. Based on a hardware architecture that uses microprocessor-controlled telephones to transmit voice over an Ethernet that also supports a voice file server and a voice synthesis server, this system has been used for applications such as directory-based call placement, call logging, call filtering, and automatic call forwarding. Voice mail, voice annotation of multi-media documents, voice editing using standard text editing techniques, and applications of synthetic voice use the Etherphones for voice transmission. Recent work has focused on the creation of a comprehensive voice system architecture, both to specify programming interfaces for custom uses of voice, and to specify the roles of different system components, so that equipment from multiple vendors could be integrated to provide sophisticated voice services.

1. Introduction

Suppose Alexander Graham Bell had waited to invent the telephone until personal workstations and distributed computing networks had been invented. What approach would he take in introducing voice communications into the modern computing environment? It was an attempt to answer this question that led to the creation of a voice communications project within the Computer Science Laboratory at Xerox PARC.

*An earlier version of this paper appeared as: D. Swinehart, D. Terry, and P. Zellweger. An experimental environment for voice system development. *IEEE Office Knowledge Engineering Newsletter 1*, 1, February 1987, 39-48.

Stated more concretely, the project's aim was to extend our existing multi-media office environment with the facilities needed to handle the transmission, storage, and manipulation of voice. We believed that it should be possible to deal with voice as easily as we can manage text, electronic mail, or images. The desired result was an integrated workstation that could satisfy nearly all of a user's communications and computing needs.

A basic requirement was to provide conventional telephone facilities (so that casual users would not have to read a manual to make a phone call), but our goals went well beyond that. We had observed that most enhanced voice communications facilities had been developed by designers of telephone systems. In contrast, we wished to draw on our experience as developers of personal information systems running on powerful workstations with graphical interfaces. We were convinced that the user would prefer to perform voice management functions using the power and convenience of workstation facilities such as on-screen menus, text editors, and comprehensive informational displays.

These aims led us to explore two related research domains:

- *"Taming the telephone"*: Despite an immense investment in research and development over the last 110 years, the user interface and the functionality of the telephone still leaves much to be desired. We contend that the personal workstation, combined with a telephone system whose characteristics we can control, make it possible to better match the behavior of the office telephone with the needs of its users. There are gains to be had in the placement of calls, the handling of incoming calls, and the capabilities available to telephone attendants (that is, switchboard operators, receptionists, and secretaries).
- *"Taming the tape recorder"*: We also believe that workstation techniques for creating, editing, and

CH2441-4/88/0000/0160$01.00 © 1988 IEEE

storing text or images can be modified to deal with digitally-recorded voice. Application areas such as electronic mail, document annotation, and dictation are candidates for improvement. Speech synthesis and recognition devices can be added to provide translation between textual and spoken information.

These two sets of activities are clearly related. A carefully designed system can support novel applications of both live and recorded voice.

In this overview we will describe the Etherphone™ system that we have developed and used to explore the integration of voice into a personal information environment. The following sections sketch the present hardware architecture, describe some of the more compelling applications that have been built to exploit it, and briefly explore the software and systems issues that have surfaced.

2. Etherphone System Description

In designing our prototype voice system, we surveyed several possible hardware architectures, including extensions of our existing Centrex service or of a commercially available PABX. Primarily because Centrex and PABX systems did not support programmable switching control, we concluded that the most effective way to satisfy our needs was to construct our own transmission and switching system[17]. Ethernet local-area networks, which provide the data communications backbone supporting personal distributed computing at Xerox PARC, have proven to be an effective medium for transmitting voice as well as data.

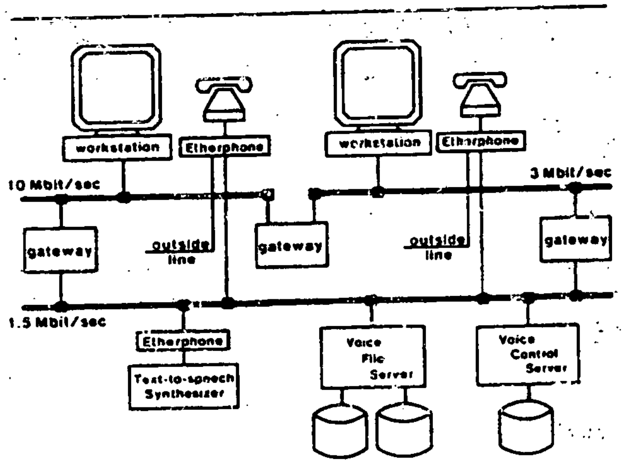


Figure 1. Etherphone system components.

Our prototype voice system consists of the following types of components connected by Ethernets, as shown in Figure 1:

Etherphones: telephone/speakerphone instruments, each including a microcomputer, encryption hardware, and an Ethernet controller. Etherphones digitize, packetize, and encrypt telephone-quality voice (64 Kbps, with silence suppression) and send it to each other directly over an Ethernet; they support conferencing by digitally summing the voice packets from other participants before converting to analog. Etherphone software is written in C. The current environment contains approximately 50 Etherphones, which are used daily by members of the Computer Science Laboratory as their only telephone service. Each Etherphone includes a connection to a standard direct-dial telephone line for access to telephones outside the Etherphone system.

We chose to separate the voice-processing functionality from the workstation for reliability, performance, and flexibility: voice-processing peripherals allowed us to provide reliable voice processing without compromising workstation performance on other tasks and to add voice to different types of workstations without changing the voice hardware. Etherphones transmit voice on a separate 1.5 Mbps Ethernet because the only Ethernet chips available to us in 1981 could not operate at standard speeds. Tests and calculations convinced us that our existing Ethernets would provide adequate bandwidth and service to carry voice as well as data. Additional information on the Etherphone hardware and the Voice Transmission Protocol can be found in a previous report[17].

Voice Control Server: a program that provides control functions similar to a conventional PABX and manages the interactions among all the other components. It runs on a dedicated server that also maintains databases for public telephone directories, Etherphone-workstation assignments, and other shared information. The Voice Control Server is programmed in the Cedar programming environment[18]. Centralized server software limited the necessary size and speed of the Etherphone processor, and thus its cost, as well as limiting the amount of C software.

Voice File Server: a service that can hold conversations with Etherphones in order to record or play back user utterances. In addition to managing stored voice in a special-purpose file system, the Voice File Server provides operations for combining, rearranging, and replacing parts of existing voice recordings to create new voice objects. For security reasons, voice remains encrypted when stored on the file server.

Text-to-speech Server: a service that receives text strings and returns the equivalent spoken text to the user's Etherphone. Each text-to-speech server is constructed by connecting a commercial text-to-speech synthesizer to a dedicated Etherphone and is available on a first-come-first-served basis. Two speech synthesizers, purchased from different manufacturers, have been installed.

Workstations: high-performance personal computers with large bitmapped displays and mouse pointing devices. Workstations are the key to providing enhanced user

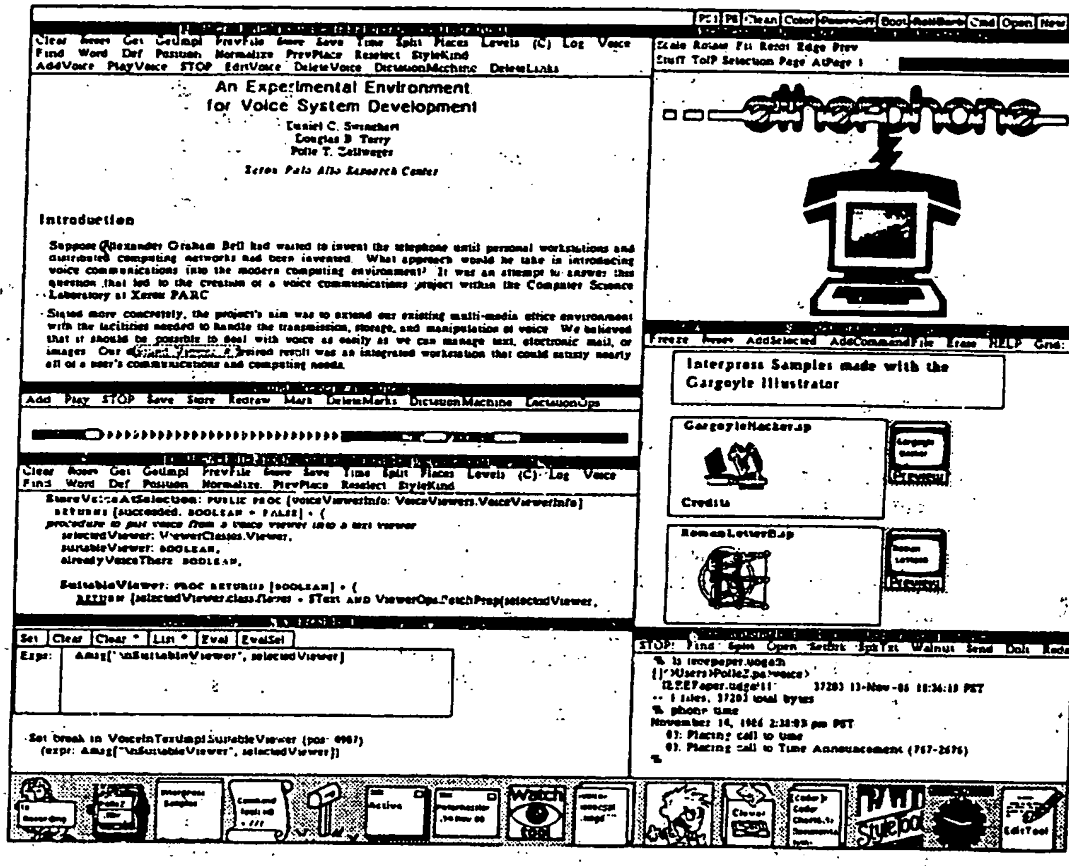


Figure 2. A Cedar screen in use. The two windows in the upper left show a document preparation task, including voice annotation of an earlier version of this paper for communication among the authors. The bottom window of this pair shows new voice (represented by arrowheads) being inserted in the middle of an existing voice annotation. The two windows in the lower left show a programming task that is monitoring part of the voice annotation system. The two windows in the upper right show images created with several graphical illustration packages. The command interpreter window at the lower right accepts user commands, similar to a Unix shell. The bottom row of icons contains files and tools that are active but are not currently being manipulated by the user.

interfaces and control over the voice capabilities. We rely on the extensibility of the local programming environment – be it Cedar, Interlisp, Unix, or whatever – to facilitate the integration of voice into workstation-based applications. Workstation program libraries implement the client programmer interface to the voice system.

In addition, the architecture allows for the inclusion of other specialized sources or sinks of voice, such as speech recognition equipment or music synthesizers.

All of the communication required for control in the voice system is accomplished via a remote procedure call (RPC) protocol[3]. For instance, conversations are established between two or more parties (Etherphones, servers, and so on) by performing remote procedure calls to the Voice Control Server. During the course of a conversation, RPC calls emanating from the Voice Control Server inform participants about various activities concerning the conversation. Active parties in a conversation use the Voice Transmission Protocol for the actual exchange of voice. Multiple implementations of the RPC mechanisms permit the integration of workstation programs and voice applications programmed in different environments.

3. Examples of Applications to Date

Most of our user-level applications to date have been created in the Cedar environment, although limited functions have been provided for Interlisp and for standalone Etherphones. This section describes the voice applications that are currently available in Cedar, including telephone management, text-to-speech synthesis, and voice annotation and editing. Figure 2 shows a typical Cedar screen using voice, text, and graphics to support programming and document preparation activities.

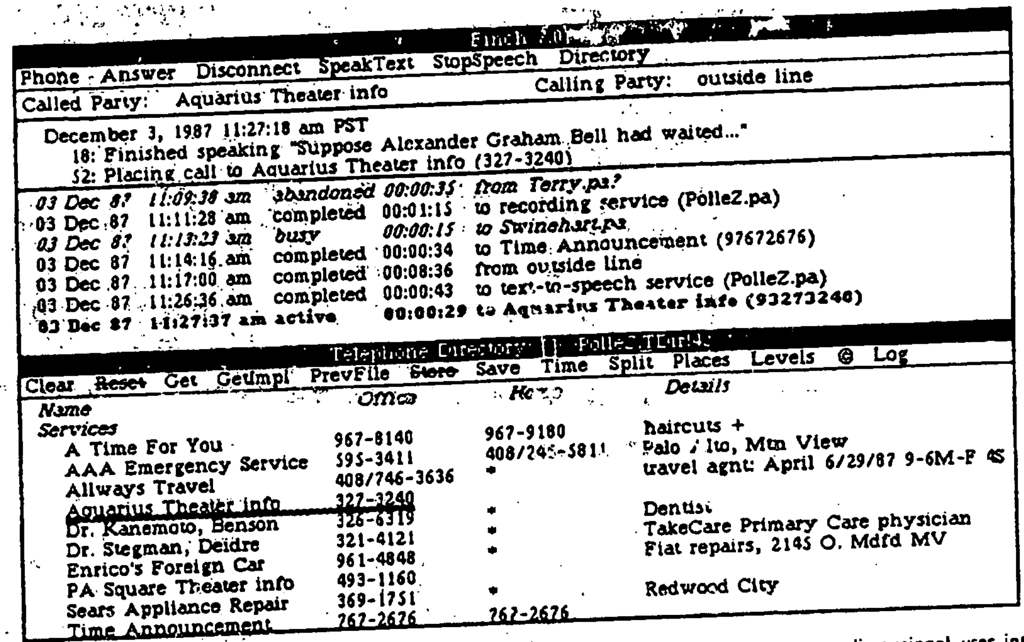


Figure 3. Two workstation telephone management windows. The upper Finch window provides a two-dimensional user interface to the Etherphone system. It includes an Etherphone control menu (the first line, including 'Phone', 'Answer', etc. buttons), a redialing area (the second line), an area for system status reports, and a conversation log (indicating a call in progress to Aquarius Theater info). The lower window shows a portion of a personal telephone directory, which is a set of speed-dialing buttons that can be created easily from an ordinary text file. The call in progress was placed by clicking on the underlined 'Aquarius Theater info' entry.

In order to make voice a first-class citizen of the Cedar environment, Etherphone functions are typically available in several ways: through an Etherphone control panel, through commands that can be issued in a command interpreter window, and through procedures that can be invoked from client programs. This integration of voice capabilities will be discussed more fully in the next section.

3.1. Telephone management

The telephone management functions provide improved capabilities for placing and receiving calls. Figure 3 shows an Etherphone control window, called *Finch*, and a personal telephone directory window.

Users can place calls by specifying a name, a number, or other attributes of the called party. A system directory database for local Xerox employees (about 1000 entries) is stored on the Voice Control Server. Users can also create personal directories, which are consulted before the system directory to locate the desired party. A soundex search mechanism[8] helps compensate for spelling uncertainties.

A variety of convenient workstation dialing methods are provided: a user can fill in fields in the Finch tool, select names or numbers from anywhere on the screen, use either of two directory tools that present browsable lists of names and associated telephone numbers as speed-dialing buttons, or redial any previously-made call by clicking on its conversation log entry. Calls can also be placed by name or number from the telephone keypad.

Calls are announced audibly, visually, and functionally. Each Etherphone user selects a personalized ring tune, such as a few bars of "Mary Had a Little Lamb". This tune is played at a destination Etherphone to announce calls to that user. The caller hears the same tune as a ringback confirmation. During ringing, the telephone icon jangles with a superimposed indication of the caller's name, as shown in the bottom portion of Figure 4. An active conversation is

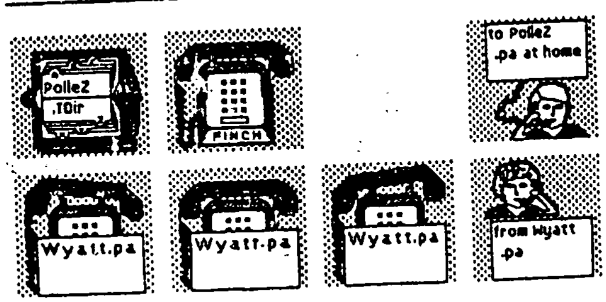


Figure 4. Etherphone system icons. The two icons at the upper left show a closed personal telephone directory and a Finch icon at rest. The icon at the upper right shows an outgoing call to Polle Zellweger's home (username PolleZ.pa). The four bottom icons show several stages of an incoming call from Doug Wyatt: the three left icons of the group are animated during ringing, while the right conversation icon is used after the call has been answered. Animation and visual feedback in the icons provide useful information without consuming valuable screen area.

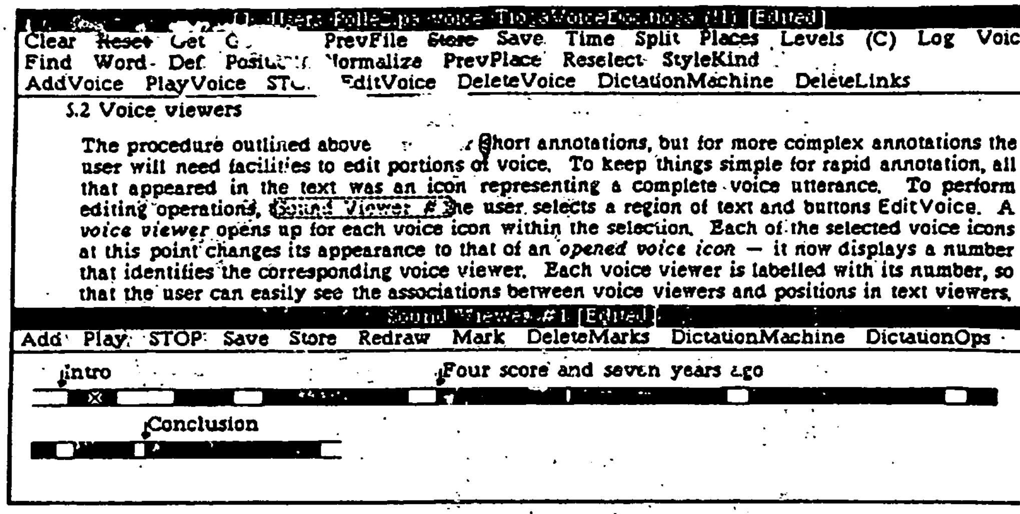


Figure 5. Voice annotation and editing. The upper window shows voice annotations being added to a Tioga document (the voice annotation system documentation). The third line of menu buttons ("AddVoice PlayVoice ...") near the top of the window are used to manipulate voice. In the second line of text, a voice annotation appears on the word "short", indicated by the comic-strip dialog 'balloon' surrounding the character. In the fifth line, a similar annotation has been opened for editing in the lower window, labelled "Sound Viewer #1". In the sound-and-silence profile in the lower window, white rectangles depict silence, while dark rectangles depict sound. The profile contains several contextual indicators to orient the user during editing. The playback cue (the gray rectangle underneath the word "score") indicates the progress of voice playback. A temporary marker in the form of a small cross has been placed in the voice section marked "Intro". The textual annotations with arrows are permanent markers that will be stored with the document.

represented as a conversation between two people with a superimposed indication of the other party's name (also shown in Figure 4). The system automatically fills in the Finch tool's Calling Party or Called Party field to allow easy redialing of the last call. It also creates a new entry in a conversation log. A user can consult the conversation log to discover who called while he was out of the office.

Our methods of locating users in an office building utilize the personalized ring tunes, which allow Etherphone users to identify calls to them wherever they may be: in their own offices, within earshot, or at other Etherphones. If an Etherphone user logs in at a workstation, his calls can be automatically forwarded to the adjacent Etherphone. An additional feature, called *visiting*, allows him to register his presence with a second workstation or Etherphone, such as during a meeting. Registering with the destination location allows users to travel more freely than forwarding calls from the home location does. Each visit request cancels any earlier requests. The common problem of forgetting to cancel forwarding is eased by ringing both Etherphones during visiting.

A more detailed series of examples that demonstrate how the Etherphone system's telephone management capabilities can increase office productivity appears in a related paper[16].

3.2. Text-to-speech synthesis

A user or program can generate speech as easily as printing a message on the display by using one of the Text-to-speech Servers. A user can select text in a display window and click the Finch tool's SpeakText menu button. A program can call a procedure with the desired text as a parameter. These features are implemented by creating a "conversation" between the user's Etherphone and a Text-to-speech Server. The system sets up a connection to the Text-to-speech Server, sends the text (via RPC), returns the digitized audio signal (via the Voice Transmission Protocol), and closes the connection when the text has been spoken. A similar mechanism is used for voice recording and playback.

Our primary uses for text-to-speech so far have been in programming environment and office automation applications. Programming environment tasks have included spoken progress indicators, prompts, and error messages. Office automation applications have included proofreading (especially comparing versions of a document when one version has no electronic form, such as proofing journal galleys) and audio reminder messages generated by calendar programs.

3.3. Voice annotation and editing

This section describes the addition of a voice annotation and editing mechanism to *Tioga*, the standard text-editing program in Cedar. More information about this system can be found in an earlier paper[2].

Tioga is a high-quality galley editor, supporting the creation of text documents that contain a variety of type faces, type styles, and paragraph formats as well as illustrations and scanned images. Tioga is the underlying editor for all textual applications in Cedar, including the electronic mail system.

the system command interpreter, and other tools that require the user to enter and manipulate text. Wherever Tioga is used, all of its formatting and multi-media facilities are available. Thus, by adding voice annotation to Tioga, we have made it available to a variety of Cedar applications. Figure 5 shows a closeup of a Tioga document containing voice annotations, one of which is being edited by the user.

Any text character within a Tioga document can be annotated with an audio recording of arbitrary length. The user interface of the voice annotation system is designed to be lightweight and easy to use, since spontaneity in adding vocal annotations is essential. Voice within a document is shown as a distinctive shape superimposed around a character, so that the document's visual layout is unaffected. Furthermore, adding voice to a document does not alter its contents as observed by other programs (such as compilers).

To add an annotation, the user simply selects the desired character within a text window and buttons AddVoice in that window's menu. Recording begins immediately, using either a hands-free microphone or the telephone handset, and continues until the user buttons STOP. A voice annotation becomes part of the document, although the voice data physically resides on the Voice File Server. Copies of the document may be stored on shared file servers or sent directly to other users as electronic mail. To listen to voice, a user selects a region containing one or more voice icons and buttons PlayVoice.

Simple voice editing is available on a visual representation of the voice. Users can select a voice annotation and open a window showing its basic sound-and-silence profile. Sounds from the same or other voice windows can be cut and pasted together using the same editing commands supported by the Tioga editor. Editing is done largely at the phrase level, representing the granularity at which we believe editing can be done with best results and least effort for an office situation*. A lightweight 'dictation facility' that uses a record/stop/backup model can be used to record and incorporate new sounds conveniently. The dictation facility can also be used when placing voice annotations directly into documents.

Sound-and-silence profiles alone do not supply adequate contextual information for users to identify desired editing locations, so several contextual aids are provided. A playback cue moves along the voice profile during playback, indicating exactly the position of the voice being heard. While playback is in progress, a user can perform edits immediately or mark locations for future edits. Simple temporary markers can be used to keep track of important boundaries during editing operations, while permanent textual markers can be used to mark significant locations within extended transcriptions. Finally, the system provides a visual indication of the voice-editing history in an editing window. Newly-added voice appears in a bright yellow color, while less-recently-added phrases become gradually darker as new editing operations occur. This use of color is similar to Lippman *et al's* use of color in text editing[9].

The voice annotation facility is implemented on top of a flexible set of voice editing and management primitives[20]. Instead of rearranging the contents of edited voice files, the Voice File Server builds a data structure to represent the edited voice and stores it in a server database. This data structure, called a *voice rope*, consists of a list of intervals within voice files. Voice is included in documents, electronic mail, and client programs solely by reference to voice ropes. To promote sharing voice throughout our distributed personal computing environment, voice ropes are immutable: recording and editing produce new voice ropes rather than modify existing ones. A modified style of reference counts enables unwanted voice ropes to be garbage collected.

4. Progress toward a Voice System Architecture

The original goals of the Etherphone project were to produce experimental prototypes that could "tame the telephone" or "tame the tape recorder" in novel and useful ways. As the project developed, however, a more fundamental goal emerged: to create and experimentally validate a comprehensive architecture for voice applications. The best way to explain the value of a voice architecture is to list some of the properties it should have:

- *Completeness*: It must be able to specify the role of telephone transmission and switching, workstations, voice file servers, and other network services in supporting all the kinds of capabilities we have identified, such as telephone services and recorded voice services.
- *Programmability*. It must permit workstation programmers to modify existing voice-related applications and to create new ones. Simple applications should be easy to write, requiring little or no detailed understanding of how the system is implemented. More elaborate applications might require a more thorough knowledge of the underlying facilities. The architecture must be designed to minimize the effect of faulty programming on the reliability and performance of the overall system. (Users of experimental software might experience program failure or reduced performance, but other users should not.)
- *Openness*. It should define the role of each major component, so that different kinds of components could be used to provide the same functions. In this way, multiple vendors of telephone and office equipment could cooperate to provide advanced voice functions in conjunction with workstation-based applications. For example, a conventional PABX (business telephone system) could be used in place of the Etherphones to provide voice switching. Ideally,

*Editing on a word or phoneme level causes two difficulties: the system cannot easily identify word or phoneme boundaries for the user, and inserting or deleting words or phonemes is likely to cause results that sound choppy.

such an architecture would evolve into a standard for voice component interconnection.

The development of the Etherphone system has included an ongoing effort to define such an architecture, and to implement the system in compliance with it. Following the general methodology employed by such modern communications architectures as the ISO reference model[7] or the Xerox Network Systems protocols[22], the voice architecture is expressed as a set of layers, each calling on the capabilities of the layer below it through well-defined interfaces or protocols. We have identified five distinct layers. From highest to lowest, these are the *Applications layer*, the *Service layer*, the *Conversation layer*, the *Transmission layer*, and the *Physical layer*.

The best way we have found to explain this organization is from the inside out, beginning with the heart of the architecture, the *Conversation layer*. It provides a uniform approach to the establishment and management of voice connections, or *conversations*, among the various services. It also provides a standard method for distributing conversation state transitions and other progress reports to the various participants in each conversation. All communications between services are mediated by Conversation layer facilities. In the Etherphone system, the functions of the Conversation layer are implemented entirely within the Voice Control Server. However, the architecture does not mandate centralized control. For example, Etherphones built with larger memories and more powerful processors could support a distributed implementation, each managing the conversations that it or its associated workstation initiated.

The *Service layer* defines the various voice-related services — such as telephone functions, voice recording and storage, voice playback, speech synthesis, and speech recognition — that form the basis for the voice applications. Each of the services must follow the uniform Conversation layer protocols in creating voice connections with other services. However, each can register with the Conversation layer additional service-specific interfaces (protocol specifications). Connections may be formed between similar services (as in a call from one telephone to another), or among different services (such as a connection from a telephone to the recording service, mediated by a workstation program). It is not expected that ordinary programmers will produce new services; the services provide both the basic user facilities and interfaces to the building blocks for higher-level applications. In the Etherphone system, some services are implemented on the server machine that contains the Voice Control Server, others on separate server machines, still others on individual workstations.

The *Applications layer* represents client applications that use the voice capabilities of the architecture. To establish voice connections, a client uses simplified facilities provided by a service that resides on the workstation along with the application. Client applications also negotiate with the Conversation layer to gain access to specialized interfaces provided by other services. The previous section illustrated many of the present voice applications using the Etherphone environment.

Logically below the Conversation layer is the *Transmission layer*. This layer represents the actual methods for representing digital voice, for transmitting and switching voice, and for communicating control information among the components of the system. In the Etherphone system, voice is transmitted digitally, in discrete packets, using a standard 64 kilobits/second voice representation and our own voice transmission protocol. Other transmission and switching methods could be substituted without affecting the nature of the programs operating in layers above the Conversation layer. Possibilities include synchronous digital transmission, or even analog transmission. The only requirement is that these components provide interfaces that allow the implementation of Conversation layer protocols. As we have mentioned, the control protocol selected for all control communications in the system was a locally-produced remote procedure call protocol. Other remote procedure protocols or message-based protocols would work equally well.

Finally, the *Physical layer* represents the actual choice of communications media, for the transmission of both voice information and control (not necessarily the same media). Besides the research Ethernet that we use (operating at 1.5 megabits/second), voice transmission on standard Ethernets, synchronous or token-oriented ring networks, digital PABX switches, or analog telephone switches could be used.

Looking at the architectural layers, it becomes easy to see how our efforts differ from work being done elsewhere. Most of the current efforts to "integrate voice and data", such as those systems built around the Integrated Systems Digital Network (ISDN) definitions[4], deal only with the Transmission and Physical layers. Other systems that include voice, such as the Diamond research effort at BBN[21] and commercial voice mail services, support some specialized applications exhibiting very scanty Service and Conversation layers. They mostly build directly on capabilities corresponding to our Transmission layer. By contrast, we have concentrated our efforts on Conversation and Service layer specifications, and on the architecture in general. A project recently initiated at Bell Communications Research has been exploring similar goals and methods[6].

To date, only one instance each of the Physical, Transmission, and Conversation layers has been implemented. We have used the resulting facilities extensively to produce the various Services and Applications described in the preceding sections. We have produced a relatively complete workstation service for Cedar workstations, and a preliminary implementation for Interlisp.

We are not yet fully satisfied with the architecture, particularly the interface between the Conversation and Service layers. This interface has proven to be somewhat clumsy to use, while at the same time restricting the number of capabilities that can be readily produced. Recent progress is encouraging, however.

5. Related Work

Incorporating voice and/or telephony into a workstation environment is an active research area.

Others have explored the use of local area networks to transport packetized voice. For example, the packet XCS system developed by Sincoskie and his colleagues at AT&T Bell Laboratories and later at Bell Communications Research also uses Ethernet[5]. In the Island system, the slotted Cambridge ring supplies guaranteed transmission bandwidth: voice and data are transported on the same network[1].

Work at IBM by Ruiz addressed both voice and telephony[12]. Off-the-shelf voice and telephone hardware was added to an office workstation. The system's prototype applications included directory-based autodialing, call logging, voice mail, telephone message recording and retrieval, and voice annotation of documents (one annotation per line). Annotated documents were stored in three parts: the text, a file of pointers to voice annotations, and a separate file for each annotation. However, there were no voice editing capabilities except for the ability to append new voice, and there was no visual representation of the recorded voice.

The Sydis Information Manager is a commercial system that provides voice and telephony[11]. It includes a shared central processor cluster with voice file storage, a voice processor, and a connection to the local PBX. Special workstations with integrated telephones called VoiceStations support directory-based autodialing, voice and text mail, dictation and voice editing, voice annotation of documents (again, one annotation per line), and voice as a user interface output medium for items such as help messages or event announcements. The Sydis voice editing system uses a sound-and-silence representation similar to ours, but it lacks the ability to annotate the voice with text markers.

Maxemchuk's speech storage system[10], developed at AT&T Bell Laboratories in the late 1970's, provided many of the same facilities for recording, editing, and playing voice as our Voice File Server. Several additional features were included to permit scanning long voice messages: text markers that could be searched by a text editor, the ability to increase playback speed, and the ability to skip forward or backward in the message. Maxemchuk's system edited voice via divide and join operations that modified the control sectors of the stored voice messages. By contrast, our scheme of building data structures that reference segments of voice files allows many users and/or voice annotations to share a single copy of the voice bits.

The Diamond multimedia message system[21], like the Etherphone system, manages documents that contain various media elements by reference. Voice annotations appear in a document as an icon with a text caption; thus their placement is even more restricted than the one-annotation-per-line systems above, but the captions are available for text searches. Because only Diamond documents can reference voice passages, they can use a simpler reference counting scheme than ours.

The MICE project at Bell Communications Research has constructed a centralized voice architecture to support rapid prototyping of communications services[6], including remote call-forwarding, voice paging, voice mail, and combined voice/data messages with editing. The central control process of the MICE system interprets finite state logic tables that specify the voice services. These tables allow changes to algorithms and services without recompiling or reloading the system (which would interrupt service to users).

In contrast to the MICE and Etherphone centralized server architectures, the approach used by Schmandt and his colleagues at the MIT Media Laboratory associates a PC-based voice peripheral with a workstation[15]. The voice peripheral uses commercial speech and telephony cards to support voice filing and editing, telephony functions, and speech recognition and synthesis. The system has been used to provide several sophisticated applications, including combined voice editing and recognition[13], conversational telephone answering[14], and office activity management.

6. Current and Future Directions

Recall that the Etherphone project began with a set of applications-level functional goals for taming the telephone and the tape recorder. After building a basic hardware platform for achieving these goals, early applications-level efforts convinced us of the need for a voice system architecture. More recent efforts have focused on developing the architecture and a few interesting applications to demonstrate and explore its unique characteristics and flexibility. Voice project members have built most of the applications, although a few programmers have included telephone management or voice synthesis activities in their applications using interfaces provided by the Services layer.

We have begun to explore a number of new directions and enhancements to the current capabilities. We have a skeletal implementation of call filtering that provides options based on the subject, urgency, or caller's identity to decrease the intrusiveness of the telephone for the callee. Our plan to integrate telephone conversation logs into the electronic mail system should have a side benefit of making the additional filtering information natural for the caller to supply.

We are working on novel kinds of interactive voice connections, such as all-day "background" telephone calls, use of the telephone system to broadcast internal talks or meetings (as a sort of giant conference telephone call), and conference calls that allow side-conversations to take place. We have recently made it possible to register conversations by name (such as BudgetMeeting or RadioService); users may join these conversations as either listeners or participants.

We plan to use our hardware-supported conferencing capabilities to incorporate text-to-speech or recorded voice into telephone calls. Among possible uses for text-to-speech are reading electronic mail over the telephone to a remote lab member as in PhoneSlave[14] or MICE[6] (but without dedicating a synthesizer solely to this task) and playing program-generated messages to callers, such as prompts or

reports of the user's location (possibly by consulting the user's calendar program, such as "Dr. Smith is at the Distributed Systems seminar now, please call back after 5 o'clock").

We are also exploring a novel *scripting* mechanism for creating viewing paths through one or more electronic documents[23]. Built on the capabilities of the voice architecture, scripted multi-media documents can contain a combination of text, pictures, audio, and action. Scripts can be used in a wide variety of ways, such as for formal demonstrations and audio-visual presentations, for informal interpersonal communications, and for organizing collections of information. Scripted documents are a dynamic form of hypermedia document whose additional structure can be layered on top of existing Tioga documents.

Finally, we would like to extend the system to other media, such as still and real-time video, other workstations, and other architectures.

In concert with work on providing new features, we intend to continue to investigate the underlying systems and theories. Managing real-time and stored voice in a distributed environment presents many interesting problems in the areas of distributed systems[19], user interface design, and voice transmission and processing technologies.

Acknowledgments

Larry Stewart, Severo Ornstein, and Susan Owicki were instrumental in the design and implementation of the Etherphone telephone management facilities. Lia Adams developed an early prototype of voice in electronic mail. Stephen Ades was responsible for the voice annotation system, and Luis Felipe Cabrera participated in the design of voice ropes and an analysis of the Voice Transmission Protocols. Dan Greene designed a prototype call filtering mechanism. John Osterhout implemented the Voice File Server, and Ken Pier implemented ring tunes.

References

1. S. Ades. *An Architecture for Integrated Services on the Local Area Network*. PhD thesis, Cambridge University, February 1987.
2. S. Ades and D. Swinehart. Voice annotation and editing in a workstation environment. *Proc. of AVIOS'86 American Voice Input Output Society Voice Applications Conference*, September 1986, 13-28. Also available as Xerox PARC report CSL-86-3, September 1986.
3. A. Birrell and B. Nelson. Implementing remote procedure call. *ACM Trans. Computer Systems 2*, 1, February 1984, 39-59.
4. M. Decina. Progress towards user access arrangements in Integrated Services Digital Networks. *IEEE Trans. Communications 30*, September 1982, 2117-2130.
5. J. DeTreville and W. D. Sincoskie. A distributed experimental communication system. *IEEE Journal on Selected Areas in Communications SAC-1*, 6, December 1983, 1070-1075.
6. G. Herman, M. Ordun, C. Riley, and L. Woodbury. The Modular Integrated Communications Environment (MICE): a system for prototyping and evaluating communications services. *Proc. of International Switching Symposium '87*, Phoenix, AZ, March 1987, 442-447.
7. International Organization for Standardization. ISO open systems interconnection – Basic reference model. ISO/TC 97/SC, *16*, 719, August 1981.
8. D. Knuth. *The Art of Computer Programming, Volume 3*. Addison-Wesley, 1973 page 392.
9. A. Lippman, W. Bender, G. Salomon, M. Saito. Color word processing. *IEEE Computer Graphics and Applications*, June 1985, 41-46.
10. N. Maxemchuk. An experimental speech storage and editing facility. *Bell System Technical Journal 59*, 8, October 1980, 1383-1395.
11. R. Nicholson. Integrating voice in the office world. *BYTE 8*, 12, December 1983, 177-184.
12. A. Ruiz. Voice and telephony applications for the office workstation. *Proc. 1st International Conference on Computer Workstations*, San Jose, CA, November 1985, 158-163.
13. C. Schmandt. The Intelligent Ear: a graphical interface to digital audio. *Proc. IEEE Conference on Cybernetics and Society*, October 1981, 393-397.
14. C. Schmandt and B. Arons. Phone Slave: A Graphical Telecommunications Interface. *Proc. Society for Information Display 1984 International Symposium*, San Francisco, CA, June 1984.
15. C. Schmandt and M. McKenna. An audio and telephone server for multi-media workstations. *Proc. 2nd IEEE Conference on Computer Workstations*, Santa Clara, CA, March 1988.
16. D. Swinehart. Telephone management in the Etherphone system. *Proc. of GlobeCom'87, IEEE Communications Society Conference*, Tokyo, Japan, November 1987, 392-402.
17. D. Swinehart, L. Stewart, and S. Ornstein. Adding voice to an office computer network. *Proc. of GlobeCom'83, IEEE Communications Society Conference*, November 1983, 392-402. Also available as Xerox PARC report CSL-83-8, February 1984.
18. D. Swinehart, P. Zellweger, R. Beach, and R. Hagmann. A structural view of the Cedar programming environment. *ACM Trans. Programming Languages and Systems 8*, 4, October 1986, 419-490.
19. D. Terry. Distributed System Support for Voice in Cedar. *Proc. of Second European SIGOPS Workshop on Distributed Systems*, August 1986.
20. D. Terry and D. Swinehart. Managing stored voice in the Etherphone System. To appear in *ACM Trans. Computer Systems 6*, 1, February 1988. An extended abstract appears in *Proc. of Eleventh ACM Symposium on Operating System Principles*, Austin TX, November 1987, 103-104.
21. R. Thomas, H. Forsdick, T. Crowley, R. Schaaf, R. Tomlinson, V. Travers, G. Robertson. Diamond: A Multimedia Message System Built Upon a Distributed Architecture. *IEEE Computer*, December 1985, 65-77.
22. Xerox Corporation. *An Internetwork Architecture*. XSIS 028112, Xerox Corporation, Stamford, Conn., December 1981.
23. P. Zellweger. Active paths through multimedia documents. To appear in *Proceedings of EP'88 International Conference on Electronic Publishing, Document Manipulation, and Typography*, Nice, France, April 1988.

Table 1. Conditions for Transmitting and Display on Communication Device

Conditions for display of calling number		Data transmitted	Data displayed on receiver (example)		
			Example 1	Example 2	Example 3
Number is passed	Private telephone	Ex.) 0451234567		0451234567	
	Public telephone	C	Public telephone	Public telephone	C
Number is blocked (both private and public telephones)		P	Blocked call	Blocked call	P
Calling from network incapable of transmission		C	Out of area	Out of area	O
Network conditions prevent transmission		S			S

number could not be advised due to the conditions such as when a mobile phonies roaming (**Table 1**). Also, as an option, a function called Number Request is provided which will respond to an incoming call in which the calling number is not advised by sending a message requesting the caller to call again and advise the calling number (**Figure 2**).

When subscribers dials "136", the Number Announce service sends a voice message advising the calling number of the latest call that was received (**Figura 3**). This service will be used by usual telephones.

Furthermore, in order to prompt the proper use of calling number information, NTT has added a rider to the agreement when the service is used, that "Guidelines related to the protection of private caller information in using caller information advice service"*2 will be respected.

Information of Calling Number

Passing and Blocking

A change to the articles in the Telephone Service Contract was approved in July 1996, under which passing the calling number to the party receiving the call was included in the basic nature of telephone service. However, since there are occasions when privacy in a telephone call must be maintained, on request from the customer, the caller is now

*2 Guidelines: Guidelines issued in November 1990 by the Ministry of Posts and Telecommunications for the purpose of obtaining proper use of private information about callers, including telephone numbers.

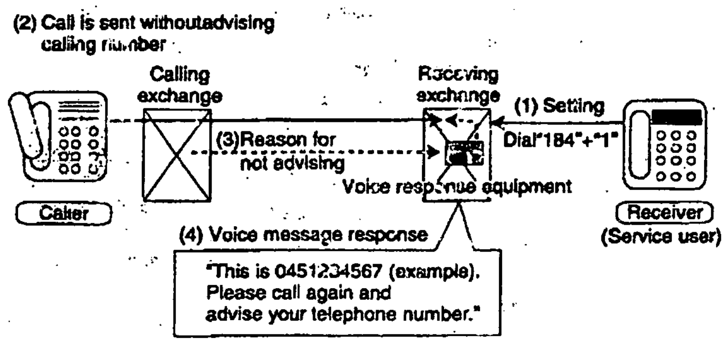


Figure 2. Summary of Number Request Service

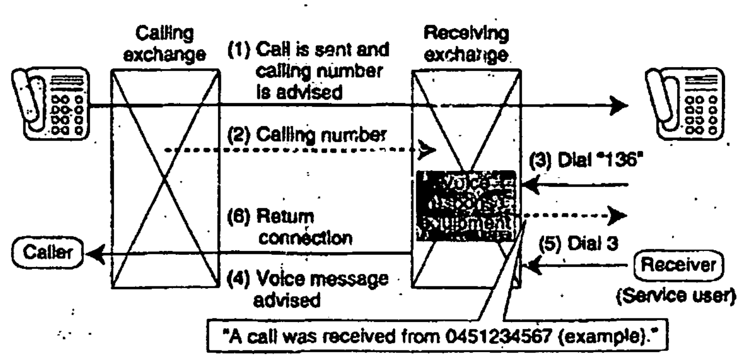


Figure 3. Summary of Number Announce Service

"Number Display" Changes the World of Telecommunications

Shun Kagata

PD 00/03/98 (6)
p. 20-25 =

This paper discusses the adoption, interconnection and principal applications (CTI-related) of services such as Number Display.

XP-000740447

Introduction

NTT began to provide a full scale "Number Display" (a service that displays the calling number) in the three areas[*1] of Yokohama, Nagoya, Fukuoka from October 1997. The service began to be offered in other areas from February 1998.

Number Display enables the subscribers to know the calling telephone number of an incoming call (hereafter referred to as the calling number) before that call is answered. The main applications for this service will be (1) as a counter to nuisance calls and to callers who will not leave a message on the answering machine, (2) for customer service work at call centers and (3) call management. A wide range of uses is envisaged in both family life and industrial activities. With a background of recent rapid popularization of the Internet and electronic mail and reduced prices of highly functional personal computers and peripheral devices, in fields such as call center related systems in particular, computer telephony integration (CTI), a move to create a new communication environment through the integration of telephone and computer, is appearing.

Summary of Services

Number Display is a service in which the caller's telephone number is displayed on a display on the receiver's telephone before the call is answered (Figure 1). This service requires a communication device with a specific function for receiving the calling number. At the time of the trial presentation, the data being advised from the network to the communication device included "Calling number", "C", "P" and "O", with the addition of "S" indicating a case in which the calling

*1 Three areas: Calling number display service had been tried in these areas from end January to end June 1997.

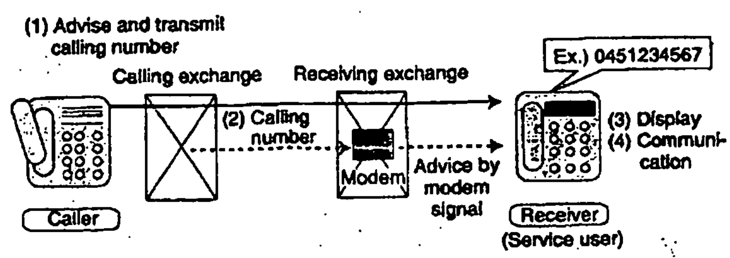


Figure 1. Summary of Number Display Service

up customer data was for a receptionist to ask the caller or to have the customer enter the information with a IVR*3 which is seems met so efficient. However, with this service, data such as a telephone number passed from the network is read directly into a computer and is used as the key for looking up customer information, which can then be displayed on the computer screen of the person receiving the call before the call is answered. This is called intelligent routing. This makes the work of reception very much more efficient.

Figure 4 shows an outline of a configuration of CTI. Interfaces between each device are in the process of being standardized (**Table 3**) and products with built-in sound boards and fax boards for unified message systems*4 and CTI servers with PBX functions for small-to-medium call centers are being offered.

With the popularization of the Internet in the United States and elsewhere, multimedia call centers (**Figure 5**) are beginning to appear that have the information providing capability of a WWW server and handle customers with Internet telephony while they watch a WWW page. Also, new services are appearing that integrate the IP network with the PSTN (Public Switched Telephone Network), such as that shown in **Figure 6**, where a user browsing the WWW via a dial-up connection can use Caller-ID*5 to take an interrupt call from a telephone connected to PSTN.

Plan to Upgrade Service

With regard to calling number display, the results of a survey in the trial areas in connection with future services have revealed a strong need for caller name-display. A review of conditions in the United States shows that Caller-ID had spread to

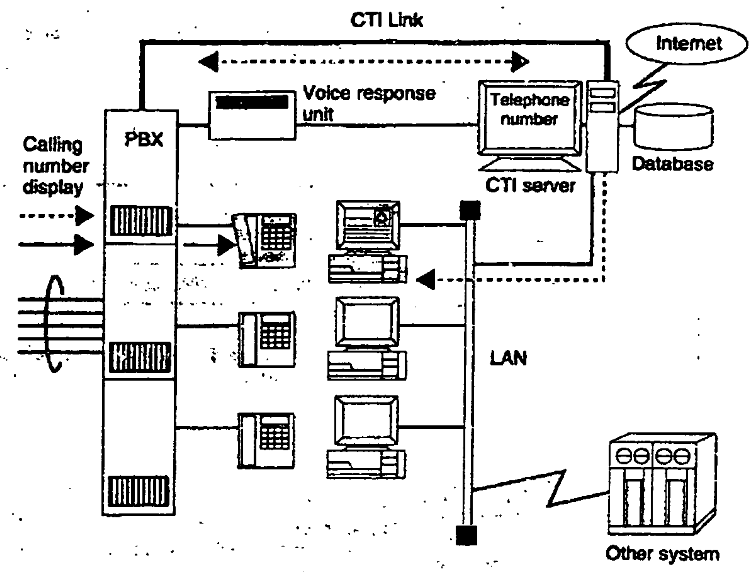


Figure 4. CTI Configuration

Table 3. CTI Standards

Protocol name	Formal title	Standardization system	Details
CSTA	Computer Supported Telecommunication Application	ECMA	Standard for PBX-computer interface
SCAI	Switch Computer Application Interface	ANSI	
TASC	Telecommunications Application for Switch and Computers	ITU-T	
TPCI	TTC-PBX-computer Application Interface	TTC	
TAPI	Telephony Application Programming Interface	Microsoft Intel	Standard for Application program interface (API) for the development of computer applications
TSAPI	Telephony Service Application Programming Interface	Novel Lucent	
SCSA	Single Computing Systems Architecture	ECTF	Standard for telephone network interface/speech processing computer boards
MVIP	Multi-Vendor Integration Protocol	GO-MVIP	

*3 IVR: An automatic telephone answering device with a voice response unit for handling caller requests.
*4 Unified message system: A message system that integrates e-mail, voice mail and faxes. Equipped with functions to announce the arrival of e-mail and voice mail to portable telephones and pagers and to convert text to voice.
*5 Caller-ID: Name of a service (calling number display service) being offered in USA from 1997.

able to control of revealing the number by using a number operation. Two types of blocking were made available by the subscriber telephone service. One of these is Per Call Blocking which involves dialing "184" before the called number to prevent the calling number being displayed to the party being called. The other is Per Line Blocking in which the display of the calling number to the party being called is blocked for all calls from a previously specified line, even when the conventional dialing procedure is followed. Per Line Blocking is provided with the Single Call Display function that cancels the blocked condition when "186" is dialed before the called number. Meanwhile, functions in the "INS Net" (Integrated Services Digital Network (ISDN) services) that display and block the calling number have been brought in line with those of the telephone service.

Changing the Calling Number

The calling number that is displayed is based on the telephone number of the line that is being used by the caller but, for convenience, the subscriber (caller) can apply for another number to be passed under certain circumstances.

If a pilot number is being used, the pilot number from a line within the same hunting group may be passed for display. If Direct Dial-In is being used, a previously specified optional number from the Direct Dial-In group may be passed. Moreover, when Free-Dial "0120" Service is being used, the Free-Dial contract line and the Free-Dial number from each line within the same hunting group may be passed. Now, the passing of Direct Dial-In and Free-Dial numbers is an extra function for which a charge is made.

With INS Net, the pilot number and the Direct Dial-In number can be passed by the calling number display function (the basic function).

Interconnection

On the premise that the caller's request with respect to revealing or blocking the number will be delivered reliably to the party being called, there is a mutual interchange of number data and other information between NTT and other carriers. When the full scale service is provided, calling numbers will be able to be passed constantly between calls within the NTT network (including calls through long distance carriers), and between most mobile and PHS carriers.

The Applications

Due the attributes of calling numbers, such as unique and widespread throughout the nation, we can think of various uses for the services in the home and in business.

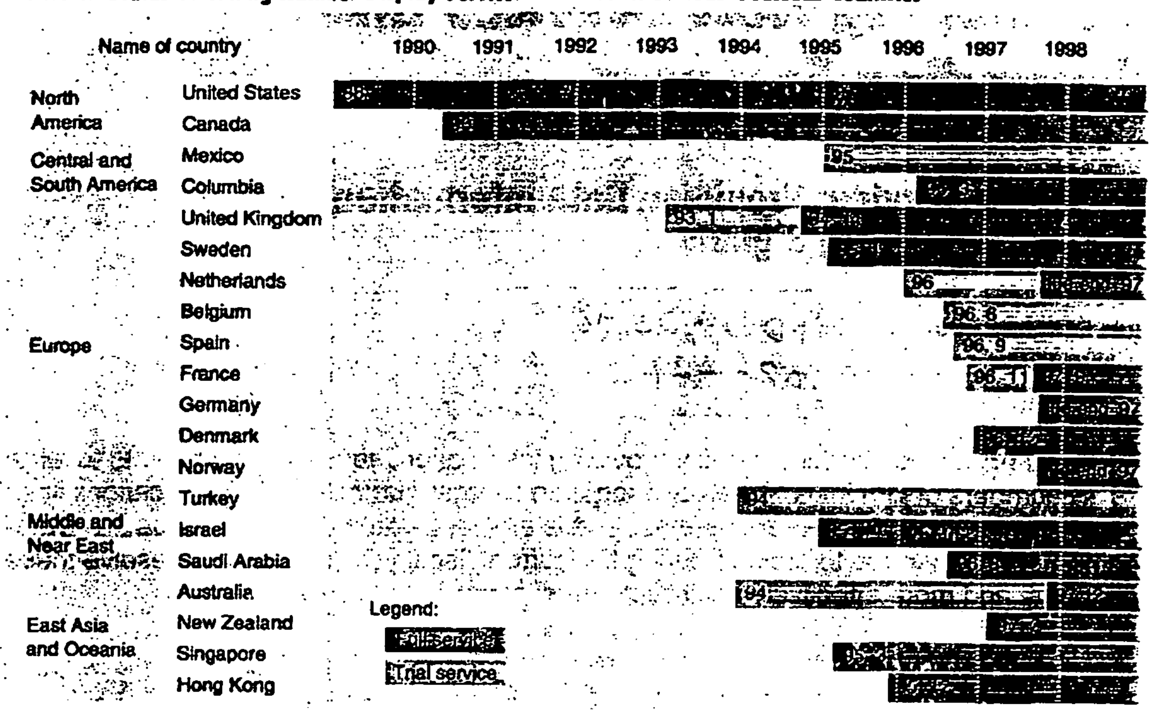
Main uses in the home could include the avoidance of nuisance calls such as non-speaking calls and persistent telephone sales calls. If the telephone number is known beforehand, the call can be handled by an answering device or the Call Refusal function (stop bell ringing). We expect that encouraging a telephone-dependent society that assumes the calling number will be transmitted, will serve to deter wilful nuisance calls. **Table 2** shows the other main functions that will be able to be used by dedicated communications devices by means of this service.

Business uses would include CTI in call centers. CTI had been advocated previously but in conventional systems, the only way to obtain information such as names and telephone numbers which are necessary for looking

Table 2. Examples of Functions Enabled by Communications Devices

Name	Function
Display calling number	Caller's telephone number is displayed
Save calling number	Caller's telephone number is saved to memory and displayed
Transmit to saved number	Initiates a call to the caller's telephone number that has been saved to memory
Display caller's name	If the caller's telephone number is the same as a previously registered telephone number, the caller's name, etc. is displayed
Distinctive ringing	Specifies a distinctive ringing sound for a call received from a previously registered telephone number and specifies a number of ringing sounds for registered telephone numbers
Display reason for number blocking	Displays the reason for the calling number not being displayed
Call forwarding	Designates of receiving equipment (cordless slave unit, etc.) in accordance with calling number
Answering method selection (Call refusal)	A call whose calling number is not displayed, or whose calling number has been registered for refusal, is answered automatically with a recorded message then treated as a nuisance call by hanging up, etc.

Table 4. Status of Calling Number Display Services Provided in Various Overseas Countries



(from hearing surveys, etc.)

numerical data) and the privacy issue.

Meanwhile, although the Number-Announce service was considered to be a service from the early period of information notification services, since it is a service that does not require specific adapters and communication devices and it is in high demand, we will continue to study it.

Conclusion

In the future, we will continue to improve customer convenience by upgrading services which meet the needs derived from new movements such as CTI.

The Author

Shun Kagata
Chief, Network Service Department, NTT.

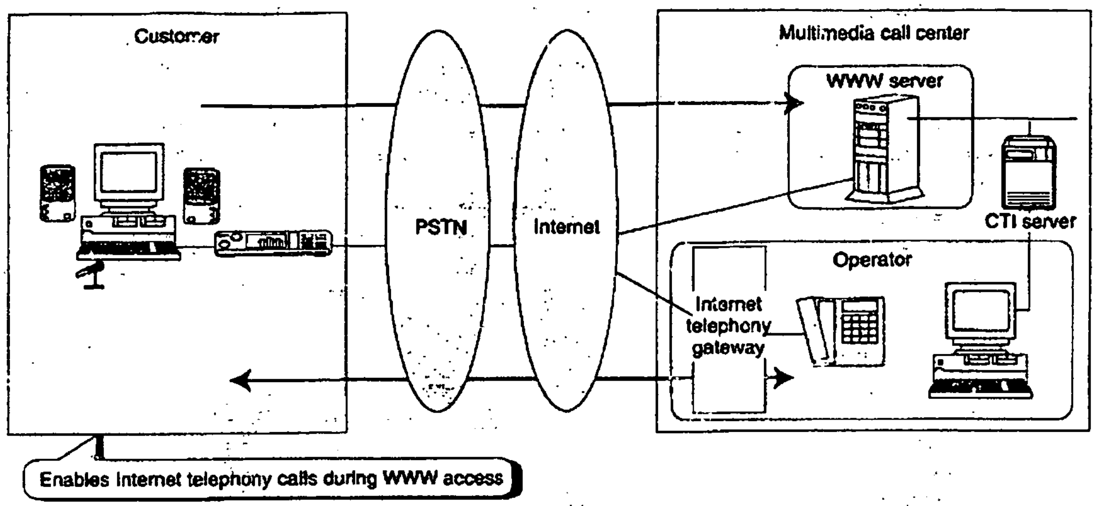


Figure 5. Outline of a Multimedia Call Center

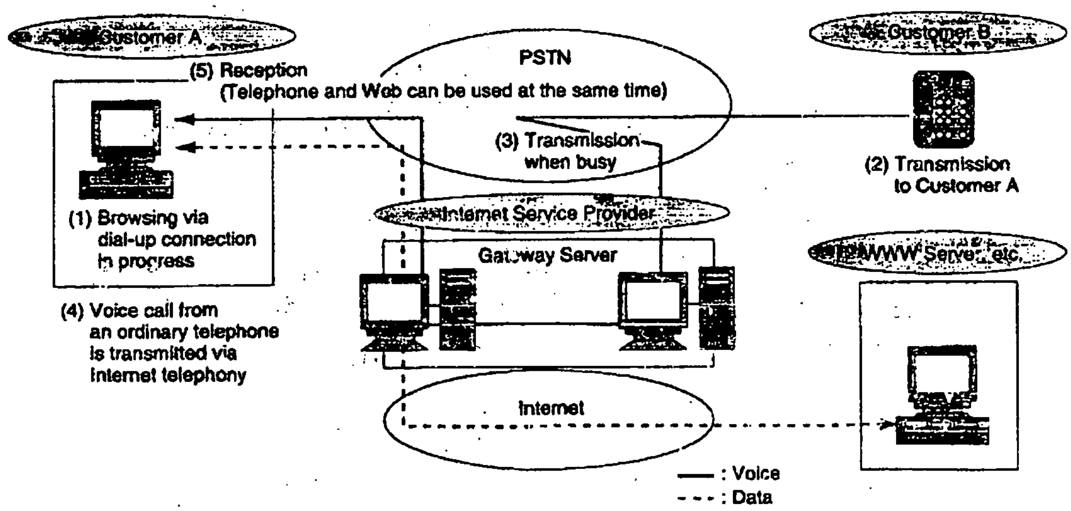


Figure 6. Example of Simultaneous Use of Internet and Telephone

about 20 % of the whole of the US at the end of 1996 and in NYNEX, about 95 % of Caller-ID seem to be shifting to Caller-Name*6. If Caller-Name is limited to within each Regional Bell Operating Company (RBOC), it can be used in almost the whole of the United States and notification of related information between RBOCs is gradually being extended as contracts for the interchange of individual names are concluded. Table 4 shows the status of services to display calling numbers in various overseas countries.

Based on this background, we would like to promote continued studies of the passing of name data and public telephone data by considering the impact on networks and communication devices (of handling other than

*6 Caller-Name: Offered as an extension of Caller-ID (for customers including in Caller-ID) which advises name of subscriber.

David Duxbury, Rob Backhouse, Mike Head, Graham Lloyd and John Pilkington

XP-000877997

Call Centres in BT UK Customer Service

E

p165–173=(9) PD. 00/10/99 VN. 8076

More and more business is being done via call centres, both within the UK and worldwide. BT has been involved in call centres for many years. This article reviews the current situation in BT UK Customer Service, and considers the next steps in this increasingly competitive and challenging environment.

Introduction

Call centres are a growing industry, in a recent Ovum Report it was estimated that by 2000 5% of the working population in Europe would be employed in call centres. It is also predicted that call-centre revenues globally will be $6 billion. Call centres have become the core of the service economy in the UK. Counting part-timers, UK call centres currently employ 1·7% of the working population, or nearly 400 000 people. And the numbers are growing. Datamonitor predicts that call-centre positions will double by 2002 before beginning to level off. Since 1994, customer calls to large organisations have roughly doubled, with call centres being largely responsible.

BT UK is a leader in call-centre technology and expertise. It operates its own call centres and manages them for other organisations. BT's own call-centre operations fall into the two broad categories of *outbound* and *inbound*. Outbound operations involve companies' staff calling out to customers, usually to offer new services, while inbound call centres handle enquiries from customers. This article looks at how the inbound call centres are presently used in the following environments, which between them are responsible for handling approximately 1 billion calls per year:

- answering Operator Assistance enquiries (100, 155 and 999 emergency services),
- giving accurate number information (192 and 153 services),
- receiving sales and billing enquiries and taking orders (150), and
- taking fault reports from residential and business customers (151 and 154).

It does not consider some of the more specialist centres (for example, Welsh language, Directory Enquiries for the blind and disabled) or inbound services and business after-sales units, although these all form an important part of the BT Customer Service call-centre management operation.

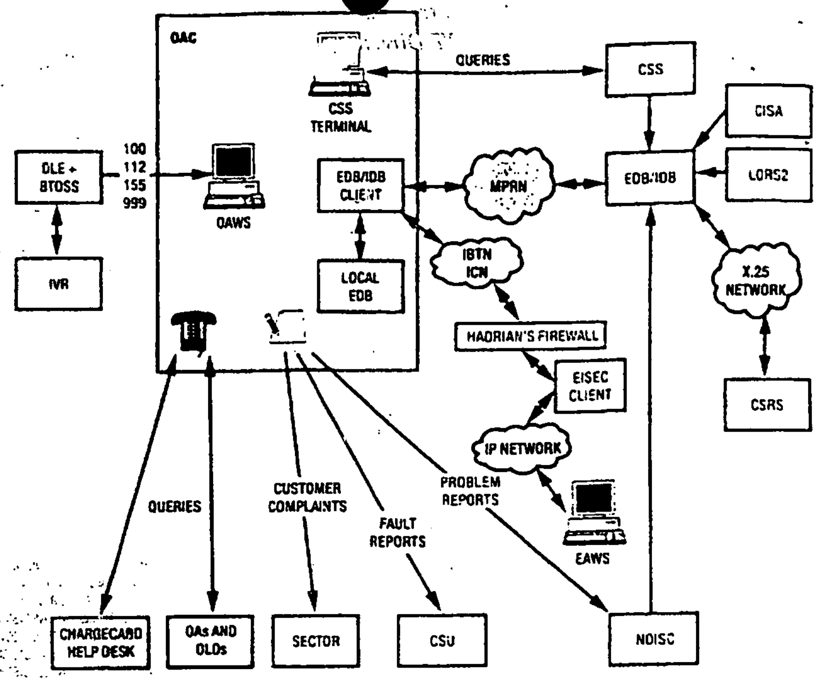
The article explores 'where we are now' and in particular focuses on systems, switching and queuing capabilities. Finally, it looks at the future of call centres and gives some indication of the way things will change.

Evolution

Although the title *call centre* would not have been used, BT has been taking calls from its customers ever since telephony was introduced. In the early days, these were large operator centres serving the geography around the exchange. Over time, separate centres for activities such as sales and fault reporting were introduced, but these again were geographically based, generally small and supported only by simple technology.

The evolution from Telephone Areas to Districts and then to a more functional organisation significantly reduced the number of centres, but there are still over 100 units and by modern standards many of these would be regarded as small. However, the introduction of both switching and systems technology has enabled BT to maximise the efficiency of these units, although more can and

Figure 1—Operator Assistance information systems



BTOSS: BT OPERATOR SERVICES SUBSYSTEM
CISA: CHARGING INFORMATION SUPPLY AND AUDIT
CSRS: CASHLESS SERVICES REPLACEMENT SYSTEM
CSS: CUSTOMER SERVICE SYSTEM
CSU: CUSTOMER SERVICES UNIT
DLE: DIGITAL LOCAL EXCHANGE
EAWS: EXTERNAL ASSISTANCE WORKSTATION
EISEC: ENHANCED INFORMATION SYSTEM FOR EMERGENCY CALLS
EDB: EXTERNAL DATABASE

IBTN: INTERNAL BACKBONE TRANSMISSION NETWORK
ICN: INTERNAL CORPORATE NETWORK
IDB: INTERNATIONAL DATABASE
IP: INTERNET PROTOCOL
IVR: INTERACTIVE VOICE RESPONSE
LORS2: OTHER LICENSED OPERATOR RECEIPT SYSTEM 2
NOISC: NATIONAL OPERATOR INFORMATION SYSTEMS CENTRE
OA: OPERATOR ASSISTANCE
OAC: OPERATOR ASSISTANCE CENTRE
OAWS: OPERATOR ASSISTANCE WORKSTATION

will be done (see paragraph on call centre futures).

Operator Assistance

Originally the operator was the only 'connect through' service for all calls. With progressive network automation, this capability is now only used when a customer requires assistance, when a customer encounters difficulty (for example, permanently engaged), or when the customer wishes to use special facilities such as alternative payment options. BT's Operator Assistance covers assistance (100, 155) and emergency (999) services and currently handles 500 000 calls per day at 15 call centres.

Network and systems

Figure 1 illustrates the information systems involved in Operator Assistance.

Common call-centre technology supports all types of call for Operator Assistance (100), emergency services (112, 999) and International Operator Assistance (155). System X BT operator services subsystem (BTOSS) switches route calls to a free operator together with associated network data—such as calling line identity (CLI) and class of service. The operator assistance workstation (OAWS), developed by BT, provides integration of the telephony and information systems, to present operators with a highly integrated and user-friendly interface to the underlying systems. This allows it to:

- obtain information from the network via BTOSS,
- obtain information from the external database (EDB) (state of line, name and address etc.),
- obtain information from the international database (IDB) (time zones call tariffs, routing, Inmarsat), and
- vet BT Chargecards using CSRS (via EDB as a gateway).

To maintain the accuracy of EDB, daily data feeds to EDB are received from both the customer service system (CSS) and other licensed operator receipt system 2 (LORS2).

The design employs object-oriented software and a script engine to achieve a very efficient and flexible user interface to capabilities which embody the required operating procedures. The workstation also incorporates voice technology, which automatically plays a greeting announcement in the operator's own voice on call arrival.

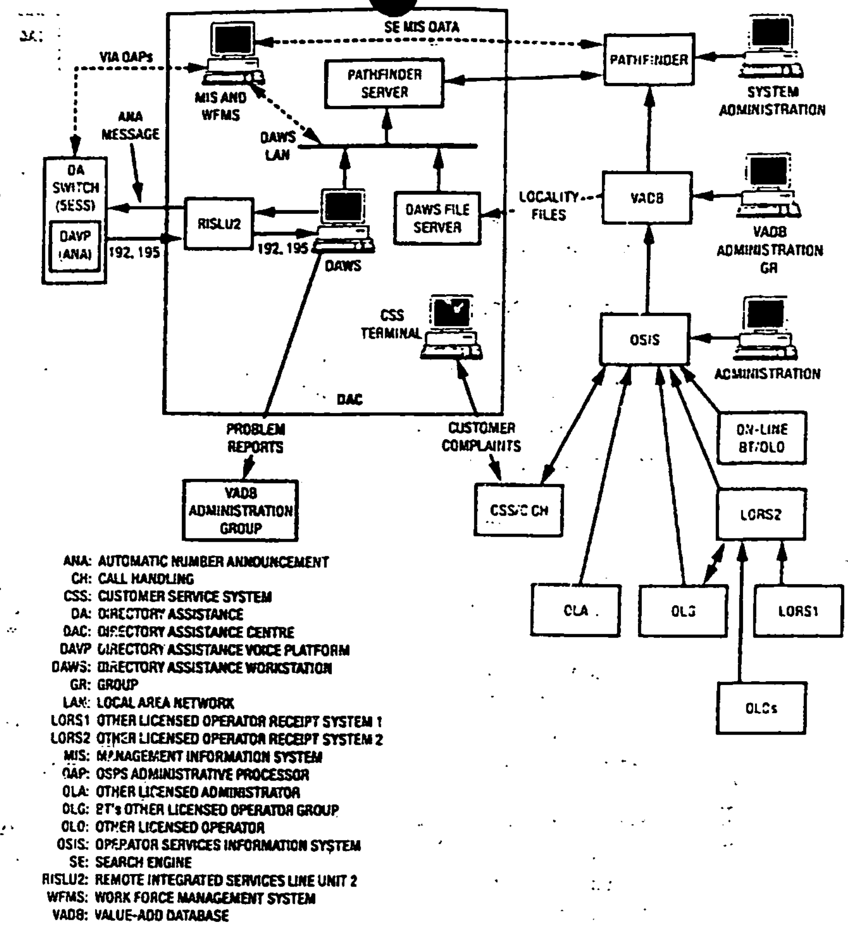
The handling of emergency (999) calls demands extremely high systems resilience and the architecture is designed to achieve 99·9999% systems availability. Information required to handle 999 calls is held locally and is available from remote servers in the event of failure. Further fall-back facilities are also provided to protect against multiple failures. Similarly, multiple levels of network fall-back are used to ensure network capacity is always available for these calls and to automatically route around network and call-centre failures. A recent enhancement has been the introduction of the enhanced information system for emergency calls (EISEC). This provides the caller name and address information for a 999 call to the emergency authority through the BT firewall, saving critical seconds in the response to an emergency.

Limited access, via the BT intranet, is also provided to the service management system. This enables operators to check for and report faults when required and transfer callers to an appropriate service department.

Directory Assistance (DA)

BT has a broad range of number information products including the well known Phonebooks and more recent innovations such as PhoneNet—the Phonebooks on the Internet (www.bt.com/phonenetuk). Directory Enquiries is a key part of the portfolio handling some 650 million calls a year at 47 DA call centres. Customer care, speed and accuracy are high priorities in handling Directory Enquiry calls. BT has recently completed the ATLAS programme, deploying state-of-the-art technology into this service.

Figure 2 – Call centre and information systems for BT Directory Enquiries



Network and systems

Figure 2 illustrates the call centre and information systems for BT Directory Enquiries.

The DA service uses Lucent Technologies 5ESS automatic call distribution (ACD) switches. Put simply, the purpose of the ACD is to place each incoming call into a single large queue so that the first free operator can answer it, irrespective of where the DA centre is located. However, for the service to operate efficiently and effectively, calls need to be handled by operators who have the best match of skills to the call. This is achieved as follows:

- Incoming calls are allocated to queues based on the locality of the caller (derived from the CLI).
- Operators are grouped into serving teams (skill groups) identified by operator identity or groups of workstations.
- Through configuration data, serving teams are allocated to each queue as the first choice to handle calls from a queue: the *primary team*. The configuration tables (Figure 3) also define the sequence of other serving teams to which a queued call is offered if a free operator in the primary team is not available within the desired answer time (the *secondary team*). The configuration tables can be changed dynamically and set-up to change automatically; for example, when a call centre opens or closes.

Figure 3 – Queue configuration table

Queue:	1	2	3	4	......	26
Primary Team:	A	B	C	D	......	Z
Secondary Team 1:	B	C	D	E	......	A
Secondary Team 2:	C	D	E	F	......	B
Secondary Team 3:	D	E	F	G	......	C
Secondary Team 4:	E	F	G	H	......	D
Secondary Team ...:	...	...	...	...	......	...
Secondary Team 25:	Z	A	B	C	......	Y

The ACDs are also integrated with management information and workforce management systems which are used to manage service quality and achieve an optimum match of operator resources to incoming call volumes.

Operator workstations

Directory Enquiry operators use the Directory Assistance workstation (DAWS). As with the Operator Assistance workstation (OAWS), this provides integration of the ACD and information systems, and employs an intelligent application. This presents operators with a highly integrated and user-friendly interface to the underlying systems and which supports the business processes. When a call arrives at the workstation, an automatic greeting in the operator's voice plays. The main function of the workstation is to provide operators with access to search the national directory information database for numbers and code information. When the correct number has been located on the database, the operator then transfers the call to the speech applications server (SAS), which automatically announces the number with the familiar phrase: 'the number you require is ...'.

Pathfinder database

The heart of the system platform is the national Directory Enquiry database (NDIS), supplied by Hewlett Packard and PCPlus. Dimensioned to hold over 150 million listings, it will perform a search and return information back to the DAWS screen in less than 0·5 seconds. The database receives 40 000 updates per night.

Customer Service and Sales, Account Services, Sales and Billing and Residential Repair

With the evolution of call centres described earlier, similar technology has also been deployed into customer service centres (CSCs) supporting 150 (account services, sales and billing) and 151 (residential repair) calls. 150 accounts for approximately

300 000 calls per day with about 70 000 per day offered to 151. However, the arrival pattern on both channels can be very irregular both during the day (see Figure 4) and through the week, with particularly high call levels being experienced on Monday morning or in periods of bad weather.

Network and systems

The 55 customer service centres throughout the UK comprise, on average, 120 agent positions and are served by two Nortel DMS-100 switches (Figure 5). The switches are located remotely from the call centres, providing ACD functions and support load balancing between the ACDs. Call routing strategies are configurable and can be evenly distributed or weighted to a particular ACD group or groups. Calls can overflow after a given threshold has been exceeded—either on number of calls in queue or length of time in queue.

SMART

As with the previous call centres an intelligent agent workstation is employed—SMART. SMART is a three-tier distributed system, which provides a graphical user interface to the billing and order entry capabilities of CSS and to the data of the products and services database (PSD). It supports a set of configurable business rules, which reduces errors and provides a simpler more efficient set of procedures compared with native use of CSS. It also provides sophisticated sales support tools which prompt advisors with on-line selling 'hints' on the products or services best suited to the particular customer.

The workstation client is a PC, hosting SMART and other applications orientated to the needs of the inbound channels. A mid-tier supports local data, particularly business rules data, a communications gateway to the CSS systems and a set of 'operability' features designed to facilitate remote management of the SMART system across all three tiers.

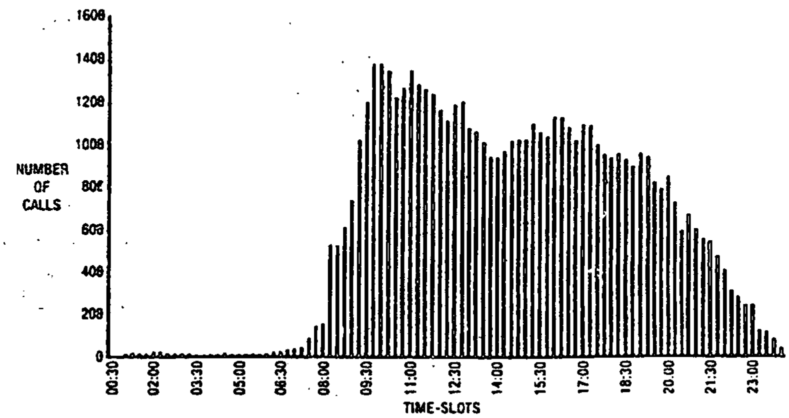


Figure 4—Typical day's call pattern into 151 repair

Management information systems (MIS)

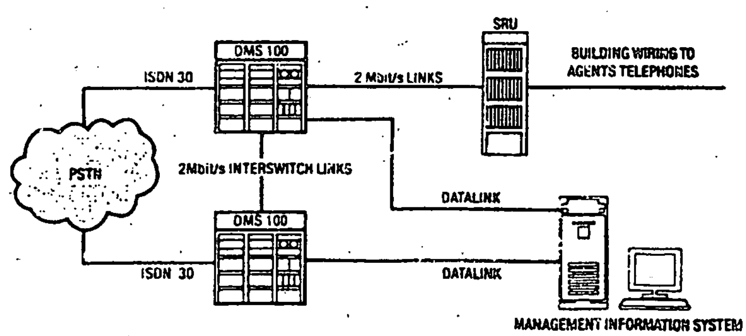
Good customer service depends on having the correct numbers of agents available to deal with the calls arriving at that particular time of the day. The resource and administration teams use management information displays to monitor service quality throughout the day. Most widely used in Customer Service is the Nortel RT1000 which displays real-time information on number of calls, number of available agents, number of calls queuing, length of time in queue, etc. Additional advisors are brought on-line when needed; however, a sophisticated forecasting and scheduling package ensures that accurate staffing is achieved based on projections derived from historical data gathered over the previous three years. RT1000 has a real-time adherence monitor, which alarms if the ACD queues are over- or under-staffed.

RT1000 also provides historical reports and is able to change the call flows in real time via load management commands to the DMS ACDs.

Call steering

In order to deal efficiently with the variety of call types being offered, and to route them to advisors with the required specialist skills, automated call steering is used. Syntellect interactive voice response (IVR) platforms are used in conjunction with Genesys computer telephony integration (CTI) servers and the Nortel DMS ACDs to capture details of the calling customer and the service request and then to route the call to the most appropriate team of service advisors. CTI is also used to present the customer account details to the advisors on call

Figure 5—DMS 100 configuration



arrival – so called *SMART screen-pops*. IVR-based call steering (Figure 6) permits a straightforward customer contact strategy while ensuring customers are routed to the correct ACD queue via simple menu selection from their telephone keypad.

Automation

Customer Service has also implemented a number of interactive customer 'self-service' applications, which enable customers to conduct simpler service transactions themselves using IVR systems. Examples include: changing Friends and Family calling circle numbers, simple billing enquiries, initial reporting and progress enquiries for faults. The fault reporting application performs a line test on the suspected faulty line and either reports the fault automatically or connects the caller to a service advisor. If the later option is chosen, the results of the line test are presented to the advisor's screen at the same time as the voice call using the CTI facilities as above. It is estimated that the automated work volume is equivalent to three call centres.

These automation services are also available to customers via the Web (bt.com) and, via the BT intranet, to advisors in other types of call centres: for example, Operator Assistance or Outbound Sales.

Business Customer Service

BT Business Customer Service has five main components comprising 154 Fault Management, 0800 666777 After Sales, Customer Relations, Simplex Management and the BT Own Use Service Centre in Gloucester. The combined units have in total around 1200 service advisors distributed across 27 centres throughout the country. The largest proportion of these people are within the 154 fault management centres (FMCs), which comprise around 600 people in 12 different locations, followed by

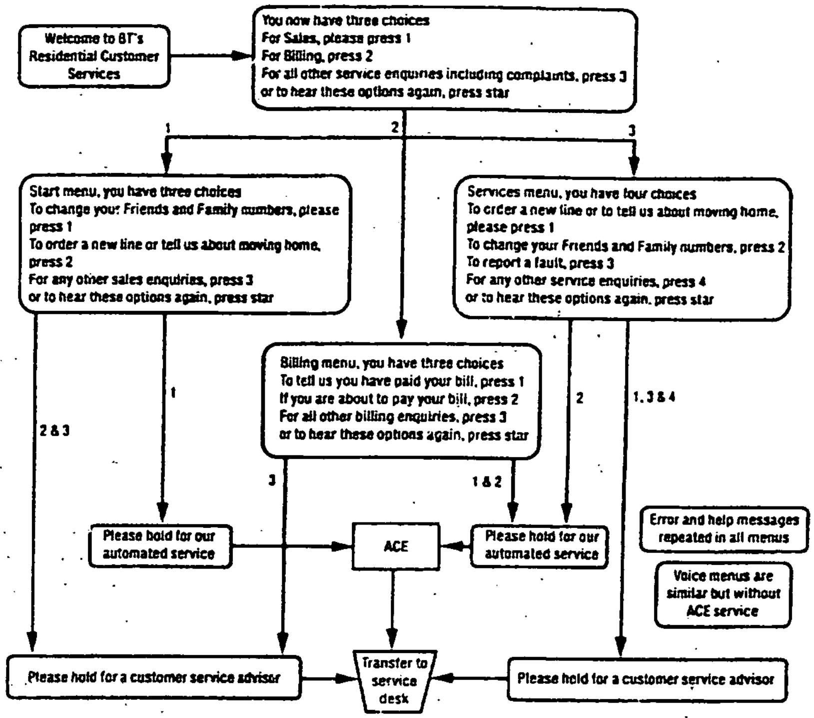


Figure 6 – 150 IVR call steering schematic summary

After Sales with around 250 people in nine locations.

The customer segments supported by these centres range from single-line business customers through small/medium enterprises up to major customers. The customer base is some 1·2 million customers with some 4 million accounts. The Fault Management and After Sales units are predominantly inbound with Customer Relations and Simplex Management involving outbound calling.

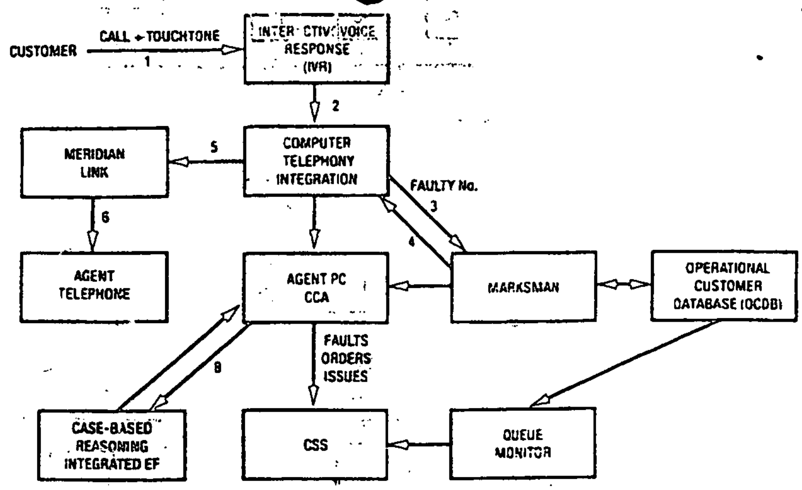
Network and systems

Compared with the residential area, the volume of calls handled are far lower but calls are much more complex. This is being fuelled by the growth in data products and services. As a result, the BT network now carries more data traffic than voice calls.

Calls are presented to advisors by customers calling 154 or via 0800 or dedicated numbers, all of which are all delivered to a network of Nortel Meridian switches: the CSC network. The CSC network then uses ACD functionality to deliver the oldest call in the network to the longest waiting agent. The CSC network also seeks to deliver calls from customers to their 'home' fault management centre: that is, if a call is identified as being generated from East Anglia, the network attempts to deliver the call to Ilford FMC first.

With the introduction of Marksman to Business Customer Service, calls are still delivered by the Meridian CSC network, but then Genesys CTI takes control of the call and the routing strategy for it. If a customer has an outstanding order, fault or issue with BT, CTI firstly tries to route the call to the advisor who originally dealt with the call. If that agent is not available, then it tries the advisor's team, then the remainder of the centre before looking for the next free advisor in the country. Planned enhancements include the introduction of intelligent call routing (ICR) enabling sophisticated call plans and the use of IVR to identify customers and the numbers

Figure 7 – Marksman overview



KEY
1 Customer rings 154, and is routed to IVR system when faulty number input using Touchtone keys.
2 Call passed to Genesys router with number attached.
3 Router accesses Marksman/OCDB using faulty number...
4 ...to retrieve customer information and agent data about who dealt with the customer last.
5 Router then routes the call to the most appropriate available agent..
6 ...via the Meridian link and switch.
7 Simultaneously the data for the call is sent via IP network to the agent's PC.
8 Agent uses CBR to resolve query and integrated EF to raise fault.

they are reporting faulty. Figure 7 shows how Marksman delivers a call using CTI, IVR and the operational customer database (OCDB).

Operational requirements

With the diversity of customers and the product set that needs supporting, speed of information retrieval is the key to resolving a problem at the first point of contact. The Marksman programme has put in place a PC and local area network (LAN) infrastructure within the service centres upon which a series of high-speed applications can be built. The applications make use of a variety of technologies, ranging from bespoke Visual Basic (VB) front-end screens to web-based information pages. The ability to test customers' apparatus is also important particularly with respect to their lines. For complex problems such as those posed by ISDN and Business Highway, Marksman uses case-based reasoning (CBR) to improve the probability of a correctly diagnosed fault resolution. Integrated with the CBR is an ISDN line test system (IFETS) which checks the customer's line end-to-end. PSTN lines can also be tested via the Marksman application. In future, line tests will be performed automatically using IVR collected data while the advisor is confirming the customer details. Further benefits this delivers are:

- More problems resolved at the first point of contact without increasing call handling times.
- The application is flexible enabling the introduction of support for new and bundled product sets
- The management information statistics (MIS) allow management of service levels and advisor performance on a real-time basis.

Figure 8 shows the components of the Marksman solution.

Marksman technical overview

The agent workstations are standard desktop PCs. Automated software delivery to the PC ensures that the current versions of software are always maintained. In addition to the Marksman application, the agent workstations are provided with standard office automation packages and e-mail. The function of the back-end systems is detailed below:

The customer handling intermediate server (CHIS) and common handling of event service system (CHESS) are mid-tier servers providing an interface between the OCDB and CSS. The mid-tier services use event notification to ensure that fault, order and issue data on CSS is updated on OCDB.

Genesys T server is made up of several different components, namely router, stat server, call concentrator, and database server. Together these components provide the heart of the CTI call delivery and drive the intranet-based statistical information on advisor performance.

The five case-based reasoning (CBR) servers provide the interface between advisors and the main CBR case base server. The five interface machines provide an equal load sharing and back-up capability for advisors logging onto case bases for resolving customer problems.

The wall-board server drives the tri-colour wall boards in the centres which are used to inform advisors of the current status regarding customers waiting or advisors waiting for a call. They also inform management of advisor status; that is, number of positions manned, number of advisors either on a call, not ready or waiting.

The script repository is used to hold scripts which automate repetitive tasks that would otherwise be time-consuming and tedious.

Call Centre Futures

One thing that can be said with certainty is that the information revolution taking place in society today will have a dramatic impact upon call centres.

BT's call centres have evolved in an age where the telephone is king, it forming the principal real time access method into the business. This philosophy is extended into both front- and back-office support systems, which today are organised around customers' telephone numbers.

The deployment and integration of desktop computing and CTI into call centres has brought about significant improvement in efficiency and the customer experience. Simultaneous voice and customer data delivery at the time of call arrival and the ability to route calls depending upon customer service advisor skills levels, are examples in use today.

This incremental approach to improving call centre infrastructure

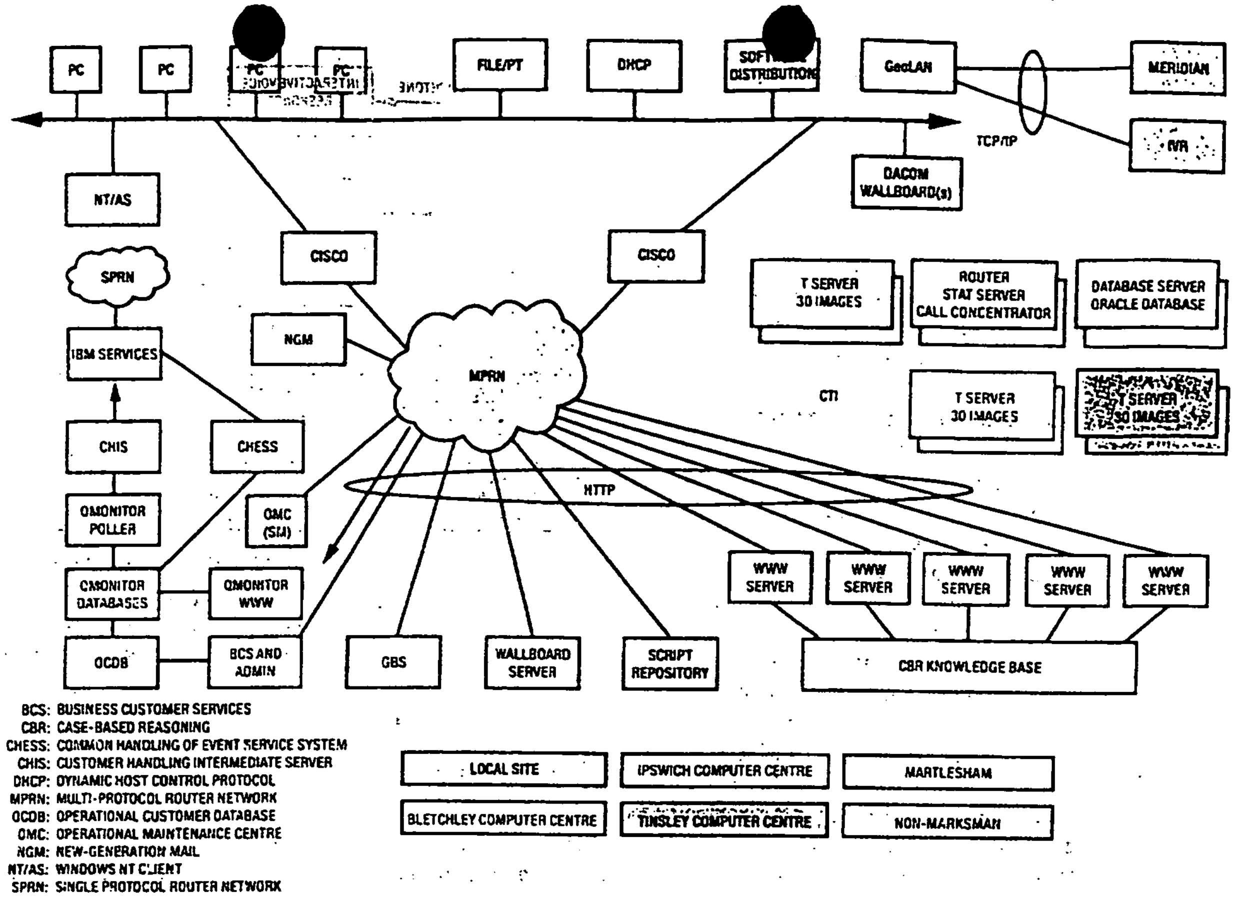


Figure 8 – Components of the Marksman solution

has kept BT in a competitive position, but is it sufficiently radical to leapfrog the competition as BT confronts the information revolution, and will it deliver the efficiencies and capabilities needed to run a world-class operation?

Site rationalisation

After extensive research, both in the UK and abroad, plans are being developed to consolidate the existing call centres. Modular centres, which are sized for the optimum number of workstations will be created to make most-effective use of accommodation and management skills. This will lead to fewer, but larger centres. These will be provided with a common systems hardware infrastructure, offering high resilience and flexibility.

Customer touch points

Today we are witnessing a massive growth in the communication channels that are being made available to customers; the Internet, mobile communications and Internet-enabled TV are all areas experiencing fast customer take up. Figure 9 illustrates the communication channels which will be available to customers.

This spectrum of communication channels will grow, as we come to terms with the transformation taking place in communications, moving from voice to a data centric environment.

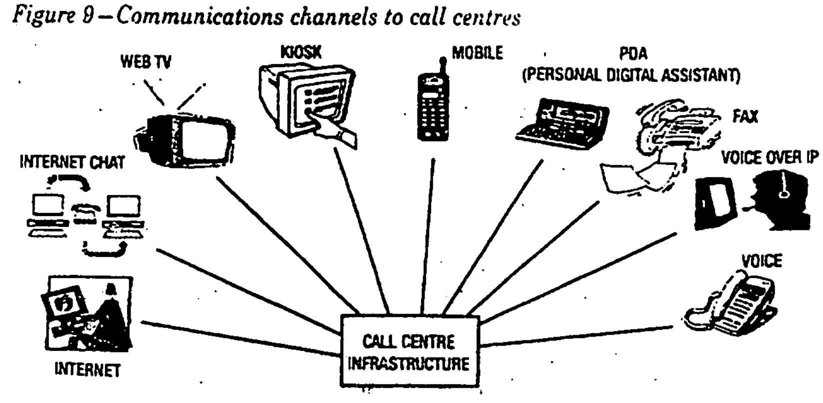
How will this change impact on BT's call centres?

Customers increasingly expect to be able to do business via the communication channel of their choice, with the 'experience' being consistent and content rich.

To satisfy this customer expectation requires a new approach to the way call centre systems are designed and delivered. Historically, these have been developed largely in isolation to satisfy the specific needs of each channel: for example, 150, 192. This has resulted in a very close fit with business requirement and world-class performance but at the expense of wide flexibility. A key constraint arising from this 'vertical' approach is to prevent the total call centre resource being regarded as a single resource 'pool' at times of peak customer demand and makes organisational change difficult.

By adopting a 'horizontal' approach to the systems architecture of call centres, both new channel integration and reuse of common

Figure 9 – Communications channels to call centres



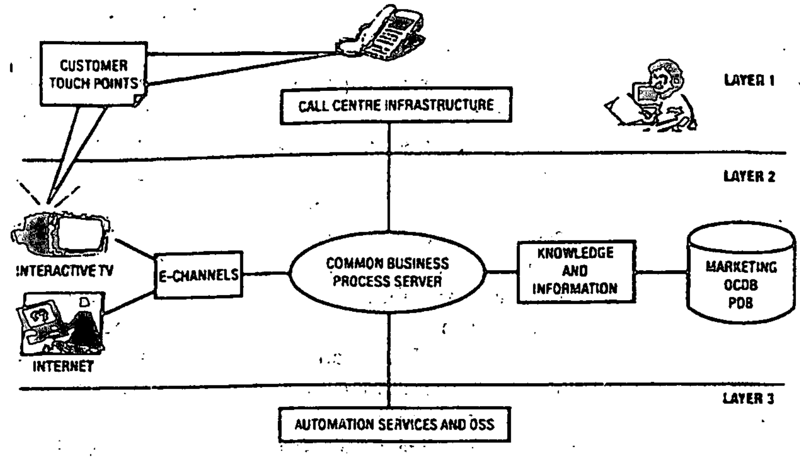


Figure 10—Three-layer model

system components can be achieved. Figure 10 illustrates the concept.

The approach has a number of benefits:

- For the construction of a new service channel, for example, Internet-enabled TV, the re-use of common business services in the layer 2 mid-tier enables both the channel and the existing call centre infrastructure to support this with minimal new development. Hence lower costs and faster time to market result when compared with the traditional approach.

- Dynamic integration at the desktop, of the applications and information required for an individual advisor's skill set increases the flexibility of the call centres and the advisor resource pool.

- Provision of integrated workflow capabilities enables a customer service advisor to work on a dynamic blend of customer service requests including voice calls, e-mail, fax, written correspondence, etc.

This infrastructure has one further major benefit: it forms the backbone by which seamless customer-centric service can be delivered, utilising a complete view of customer relationships to ensure service is tailored to their individual needs and preferences. The call centres will increasingly take on the roles of help desks, providing immediate context sensitive support to customers, with the majority of service transactions handled in a 'zero-touch' manner. In other words, service will be delivered to the customer without any person being involved except perhaps a customer engineer. In this environment, if the call centre is contacted, it will be increasingly in a secondary 'assistance' role rather than a primary 'transaction' role.

Transactions to relationships

In a business environment, where BT's core network is no longer a competitive advantage and where it is possible to build a new telco from scratch in under five months, how does BT compete?

Jack Welch, CEO of General Electric, made the following statement, 'We have only two sources of competitive advantage:

- The ability to learn more about our customers than the competition,

- The ability to turn that learning into action faster than the competition.'

BT is confronted with exactly the same challenge, a customer that has brought a product via bt.com will expect BT to know about it when they ring in with a service call that same day, and they will not expect to receive a call asking if they would like to purchase the very same product!

As the number of customer communication channels increases and BT moves towards the bundling of products via propositions, the company needs an integrated information service that supports a complete view of the customer relationship, delivered via a powerful interface to the customer service advisors. This would enable the complete customer relationship to be visible no matter how the customer makes contact.

Delivering systems and applications that support this level of functionality will be a significant challenge. However, the increasing functionality and maturity of proprietary CRM and ERP packages, means that it is feasible to configure standard software packages to achieve this rather than embark on bespoke development. Availability of 'out-of-the-box' solutions will be essential to support the evident rate of change in the marketplace.

Several initiatives are underway looking at how BT can satisfy this requirement, with vendors of both front- and back-office applications being assessed.

In summary, BT's Call Centres will change dramatically over the next few years as customer service, BT's true competitive advantage, is optimised.

Biographies

David Duxbury
Customer Service,
BT UK

David Duxbury joined the British Post Office in 1961 as an apprentice in the Leeds Telephone Area. He was later involved with exchange design in the North East Region before moving into Engineering Management Services within headquarters. He later became involved in the setting up of Account Management and Technical Support before moving to Newcastle where, with the formation of Districts, he

became Deputy District General Manager for the North East. With Project Sovereign he moved south and joined the Personal Communications Marketing Division and went on to become the Director of Retail, Consumer Products, and Cable Television Services in the old Consumer Division. On 1 October 1998 he took on the role of General Manager Call Centre Management in Customer Service. David is a Chartered Engineer and a member of the IEE.

Rob Backhouse
Customer Service
BT UK

Rob Backhouse joined BT in 1972 as an apprentice. He progressed through roles as linesman, PBX and telephone exchange engineer, PBX and network installer, project coordinator, and project manager. He was promoted to manager in 1988 in Colchester, responsible for East Anglia Districts internal PBX and ACD systems. He then joined a national group performing a similar role. He led a team which implemented Meridian 1 ACD systems in the customer service call centres throughout the Home Counties. In 1993, he moved his job to Personal Communications division in London and was soon running the project to replace the Meridian ACDs with DMS-100. He has represented BT at an international conference in Florida, USA, which involved presenting a paper describing the use of the DMS-100 facilities most efficiently in supporting networked ACD solutions. He has also assessed the suitability and specified call centre solutions for BT joint ventures in Italy and Spain.

Mike Head
Customer Service,
BT UK

Mike Head joined the General Post Office in 1966 as an apprentice in London working on repair and calibration of equipment used in repeater stations and transmission centres. Over the subsequent years, he worked on a variety of products, projects and programmes varying from loudspeaking telephone and noise-cancelling microphone design through to systems implementation. He spent four years working with Government National Accounts as a presales engineer selling computing solutions to the army and navy before moving to BT Laboratories in Martlesham to run a computer support and in-life software development group. He currently works in BT UK Customer Service as the Marksman programme manager delivering a culture change programme to Business Fault Management and After Sales.

Graham Lloyd
Customer Service,
BT UK

Graham Lloyd joined the then British Post Office in 1972 and progressed through the three year engineering apprenticeship scheme, taking up a Technical Officer post in business systems planning at Cambridge. He completed a number of major projects including the design of a replacement ISDX switch for Addenbrookes Hospital. After promotion to a managerial grade in 1988, he was given responsibility for the voice communications needs for East Anglia's 8000 workforce. Key achievements included upgrading the District voice network and designing a portfolio handbook with product and service descriptions for 800 managers. In 1991, after further promotion, he played an active role within the London Front Office project, being responsible for the design of voice systems technology utilised by Customer Service. He acted as a single interface with Nortel and designed the DMS100 network to support a single virtual queue of 950 advisors. The configuration was the first application of its kind in Europe, and the largest single node ACD in the world. In 1996, he was selected to set up a voice systems team to deliver infrastructure for the telemarketing expansion project, which subsequently achieved the fastest ramp up of call centre capacity in the history of BT—3300 advisor positions in 14 months. He is currently working in the Business Development and Operations team working upon solutions design approval process and influencing the technical direction of BT Customer Sevice's call centres.

John Pilkington
Customer Service,
BT UK

John Pilkington graduated from Leicester Polytechnic with a joint honours degree in Physics and Computing. He later went on to gain a Diploma in Management Studies from Leeds University. He joined BT Laboratories in 1987 and spent several years working on international standards where he led various CCITT sub-committees developing the X.400 messaging and X.500 directories protocols. In 1992, he moved to BT's Operator Services Directorate, whose headquarters are in Leeds. In 1997, he was asked to create a new team responsible for the requirements capture, design and integration of a platform that would replace the DA service before the millennium. With the DA programme successfully deployed on time, John's role was recently expanded. He currently leads a multi-discipline team responsible for the end-to-end management of all hardware and software releases onto the DA, Operator Assistance and 999 platforms.

Enquiries about this article should be directed to Lyn Windmill at lyn.windmill@bt.com.